THE INTERPRETATION OF LATERALITY

I.C. McManus

(Department of Psychiatry, St. Mary's Hospital, London, and Department of
Psychology, Bedford College, London)

"I can only extend my sympathy to the psychologist of the future, for it seems
as if he must first be a mathematician, then a statistician, ... and if he is not dead of
old age by then, a psychologist."

(Dunlap, 1938)

There is now, in the literature, a vast number of reports of experiments
in which a functional asymmetry has been found. This result is so secure
that it can hardly be disputed. But in most of these reports there is the
problem that, having found one's asymmetry, what is one to do with it? A
typical approach is to ask questions such as whether the asymmetry relates
to handedness, familial sinistrality or sex? And in most cases these hypo-
theses will be tested by means of an analysis of variance.

In the present paper I wish to suggest that the use of the analysis of
variance for studying laterality is inappropriate, and may lead to erron-
eous conclusions. Instead, far more explicit models must be fitted, and
their relative merits examined by means of maximum likelihood techni-
ques. In particular, without such explicit statistical models it is not pos-
sible to make such statements as, "Left-handers are less lateralised for a
task than are right-handers"; and indeed such simple statements contain
such a powerful ambiguity that they are dangerously misleading.

The nature of the problem may be shown simply by considering the
results of a hypothetical experiment. Twenty left-handers (L) and twenty
right-handers (R) carry out a simple dichotic listening experiment, and for
each subject an asymmetry score is computed.

Figure 1 shows simulated data from such an experiment the data being
generated randomly, subject to contraints to be described. Figure 1a shows
the raw data from such an "experiment". The mean asymmetry scores for
the R and L groups are .85 and .21, and the standard deviations are .77 and
1.29 (figure 1b). A t-test (had equal variances been assumed) suggests that
the difference in the means is significant ($t = 1.917$; d.f. $= 38$; $p = 0.063$,
two tailed), and an F-test suggests that the difference in the variances is
also significant ($F = 2.85$; d.f. $= 19, 19$; $p = 0.027$, two tailed). The
conventional conclusions from such an experiment would be that i) left-
handers are less lateralised than right-handers and ii) left-handers are
more variable in their lateralisation than right-handers. Neither of these

conclusions is in fact appropriate for the data of Figure 1, since they were actually drawn randomly from the compound distributions shown in Figure 1c; the distributions, for both R and L, are each made up of normal distributions with the same absolute mean (+1 or −1) and standard deviation (0.5) in each case, the only difference being that for the R group 0.1, and for the L group 0.3, of the population come from the smaller, minor distribution. The groups truly differ only in their balance of *direction* of lateralisation, not their *degree* of lateralisation within sub-groups, but the analysis of variance method used has confounded these two items to produce results which are totally misleading, and are open to a very different theoretical interpretation. It is also worth noting that by juggling with the parameters of the distributions we could have produced a situ-

Fig. 1 — a - *Results of a simulated experiment, the top line indicating asymmetry scores for a group of right-handers, and the bottom line scores for left-handers, each point representing a single subject. b - Fitted normal distribution for each set of subjects. c - Compound distributions from which the data points were actually drawn.*
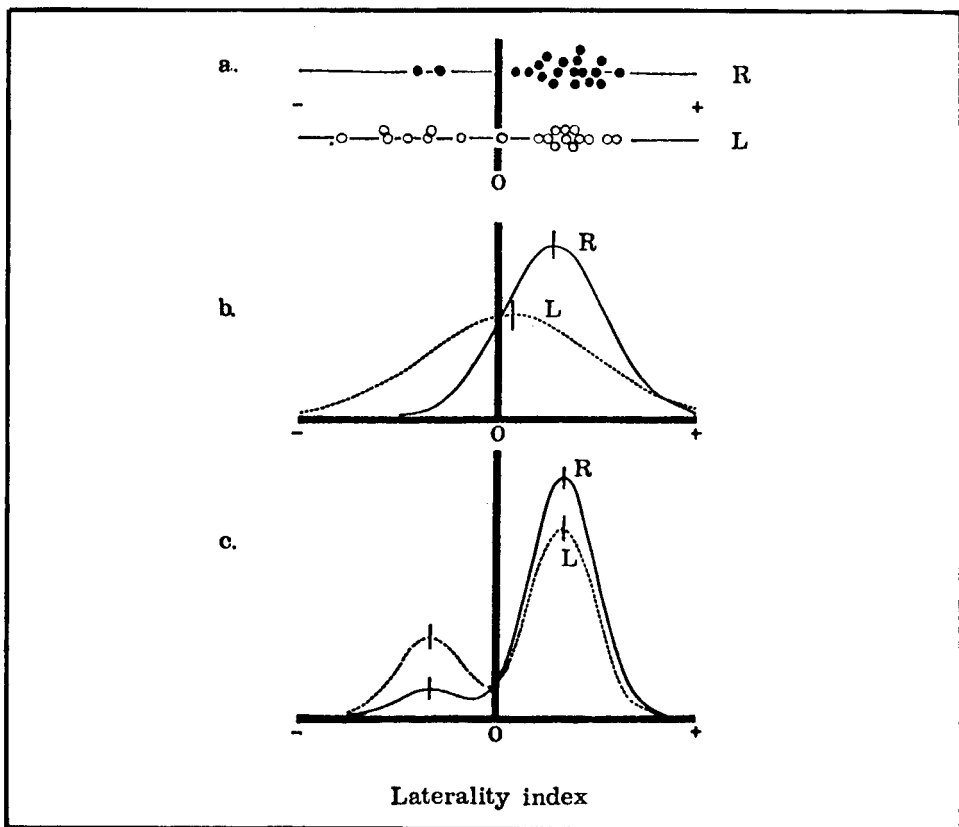
Laterality index

ation in which despite there being no difference at all in the means for the R and L groups there would truly have been differences in both direction, and degree of laterality in the distributions. Hence one may obtain both false positive and false negative results by using the analysis of variance in laterality experiments.

To be fair, it is not the analysis of variance per se which is responsible for this confounding of direction and degree, but its inappropriate application. Analysis of variance is not a neutral statistical technique into which one may put any data for a general screening for effects, but it is instead the fitting of a particular theoretical model, the theoretical model being implicit within the method but being rarely stated explicitly — and this model has several assumptions. The major assumptions of importance are that the between subject error distributions are simple normal distributions with the same variance for all sub-groups, and that all experimental effects act as additive constants. The error assumption is erroneous in the case of laterality.

It might be objected at this stage, for the example quoted above, that the bimodality of the distribution would have been noted, and a different statistical approach used. Indeed it might have been; but only if the variance within either component of the distribution was relatively small (e.g. a standard deviation of, say, 0.1). But in most laterality experiments the variance is much greater than this and hence the data from the minority distribution will tend to be hidden in the tail of the majority distribution, which will tend anyway to have some cases which are on the opposite side of the zero line. The problem will be compounded if one has a complex experimental design with several balance and design variables in addition to the experimental factors of interest. And of course the problem will not be noticed at all if the data are simply 'plugged in' to one of the large statistical packages on computers without a close examination of the raw data, the means alone being examined at the end of the analysis.

One solution to the problem might be to use the absolute asymmetry score as the dependent variable in an analysis of variance, and to include direction of asymmetry as an independent variable in the analysis. This has several problems. Firstly, the design will usually be unbalanced since there will be more +ve individuals than −ve individuals, and hence a classical experimental analysis of variance could not be used. Secondly, and more seriously, the dependent variable, the absolute asymmetry score, will be truncated, since it will be unable to be less than zero, and hence as the variance increases relative to the mean, observations will 'cross over' the zero line, and be coded as being of the other sub-type. A third problem is that even if the other objections do not apply, the analysis of variance model will still assume that the variance within all sub-groups

is identical; and yet one of the specific problems of interest in laterality is whether left-handers are more variable in their laterality than are right-handers. An analysis of variance approach could only ask whether they are, on average, more or less lateralised.

A further possible solution would be not to use asymmetry scores *per se* but instead to use the raw scores from each ear or eye and to include eye or ear (right or left) in the analysis as another independent variable. This however removes all the benefits of having an asymmetry score with pre-determined characteristics and replaces it, of necessity, with a simple difference score, with all of its known disadvantages. Furthermore the method *still* assumes that all error scores are unimodally normally distributed, and provides no means for analysing differences in variance between sub-groups.

Finally it is worth noting that none of the above problems would be solved by using non-parametric statistics (which depend merely upon the ordinal relations of the numbers) and indeed many of them would be rendered insoluble since many hypotheses are, of their very nature, distributional in type.

A simple model of population data

Consider a group of pure + ve individuals. Let their asymmetry scores have a mean of μ_+ and be distributed normally with a standard deviation σ_+. (I will assume that all such distributions are normal for the present exposition — there is no reason, in principle, why other distributions, say skewed normal, or exponential, should not be incorporated if it were felt to be theoretically necessary).

Let a proportion π of the population be of type $-$. Let the other $(1-\pi)$ be of type $+$. Let the $-$ve individuals have a distribution with mean μ_- and standard deviation σ_-. Considering just the total data from an experiment (independently of any sub-groupings due to handedness, etc.) we now have a model with five parameters which we can fit to our observed data. The simplest hypothesis, outlined earlier, may be called Model I, and says that + ve and $-$ve types are entirely symmetric except for direction, that is $\mu_+ = -\mu_-$ and $\sigma_+ = \sigma_-$. This we may regard as a base line from which to examine more complex models. If the individuals of type $-$ are truly less lateralised than those of type + then we would find increased support for a model in which, unlike Model I, $\mu_+ \neq -\mu_-$. As an alternative it may be that individuals of type $-$ are more variable in their laterality than those of type + i.e. $\sigma_- > \sigma_+$. And of course it may be that type $-$ is both less lateralised *and* more variable than the + type; that is, $\mu_+ \neq -\mu_-$ and $\sigma_- > \sigma_+$.

Let there be a total of n independent observations, one per subject, which we may represent as x_i, where $i = 1...n$. For a particular observation x_i we may calculate the relative likelihood, L_+, that the observation is derived from a normal distribution with mean μ_+ and SD of σ_+.

$$L_+(x_i) = N(x_i, \mu_+, \sigma_+)$$

$$\text{where } N(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e\,[-(x-\mu)^2/2\sigma^2]$$

... 1

And similarly the likelihood, L_-, that x_i is from a distribution with mean μ_- and SD σ_- is :

$$L_-(x_i) = N(x_i, \mu_-, \sigma_-)$$

... 2

If the types $+$ and $-$ are represented in the relative proportions $(1 - \pi)$ and π, then the overall likelihood of a particular observation x_i, given a set of parameters μ_+, μ_-, σ_+, σ_-, and π, is:

$$L(x_i) = (1 - \pi) \cdot N(x_i, \mu_+, \sigma_+) + \pi N(x_i, \mu_-, \sigma_-)$$

... 3

By multiplying the likelihoods of all the observations, x_i, we may determine, L_T, the likelihood of a particular set of model parameters given the whole set of data points:

$$L_T(x_i, i = 1, n) = \prod_{i=1, n} [L(x_i)]$$

... 4

In general it will be more convenient to derive the support rather than the likelihood, where $S(x_i)$ represents the support for a particular model, given a particular observation, x_i,

$$S(x_i) = \log_e [L(x_i)]$$

... 5

and hence S_T, the support for a particular set of parameters given all of the data points is:

$$S_T(x_i, i=1, n) = \sum_{i=1, n} S(x_i) = \sum_{i=1, n} \log_e\{(1-\pi)\cdot N(x_i, \mu_+, \sigma_+) + \pi \cdot N[x_i, -(\mu_-), \sigma]\}$$

... 6

For any particular set of parameters we may thus derive a value of S_T for their fit to a particular data set. The problem is to find the values of the five parameters which maximise the support; that is, to find the maximum likelihood estimates of μ_+, σ_+, μ_-, σ_-, and π. This problem seems to be peculiarly intractable analytically, having first been studied by Pearson (1894) and subsequently by Cohen (1967), Bhattacharya (1967) and Has-

selblad (1966), and no analytic solution to the general case seems possible at present except in the trivial case when $\pi = 0$. A very cumbersome analytic solution is available in the two group case; for the three-group case (i.e. a compound distribution composed of three normal distributions) or for the more complex cases to be described later, no analytic solutions are as yet available. Fortunately this is not a major set-back since the ready availability of modern computers means that one may use numerical methods of 'hill-climbing' to solve the problem. The present author has successfully used the subroutine E04JBF of the Numerical Algorithms Group (NAG) library (1981) to estimate solutions to such equations. The programme uses a quasi-Newtonian method whereby evaluations of the function at the present estimates of the parameters and at small steps from those estimates are used to estimate the first and second derivatives at that point, and hence to calculate a new maximum for the function and then new estimates of the parameters. The program proceeds iteratively, inserting the new estimates of the parameters and repeating the process.

Given the ability to find estimates of the parameters in equation 6 we may now use this ability in order to find improved methods of analysing such problems. We can fit a series of models to our population data:

(a) *Model 0: a unimodal normal distribution*

Let $\pi = 0$, and estimate μ_+ and σ_+ from the data. Let the maximum support be S_0.

(b) *Model I: a symmetric bimodal normal distribution*

Let $\pi \neq 0$, $\mu_+ = -\mu_-$, and $\sigma_{+,} = \sigma_-$, and the maximum support be S_I.

(c) *Model II: A bimodal distribution with different means but equal variances*

Let $\pi \neq 0$, $\mu_+ \neq -\mu_-$, and $\sigma_+ = \sigma_-$, and the maximum support be S_{II}.

(d) *Model III: A bimodal distribution with different variances but equal means*

Let $\pi \neq 0$, $\mu_+ = -\mu_-, \sigma_+ \neq \sigma_-$, and the maximum support be S_{III}.

(e) *Model IV: A bimodal distribution with unequal means and unequal variances*

Let $\pi \neq 0$, $\mu_+ \neq -\mu_-, \sigma_+ \neq \sigma_-$, and the maximum support be S_{IV} (see Figure 2).

We may now ask whether, given a particular set of data, there is sufficient evidence (or support) to reject the hypothesis that the data are derived from a unimodal normal distribution (i.e. Model 0). Since support is additive we may calculate $S_{IV} - S_0$, which is a measure of the support for
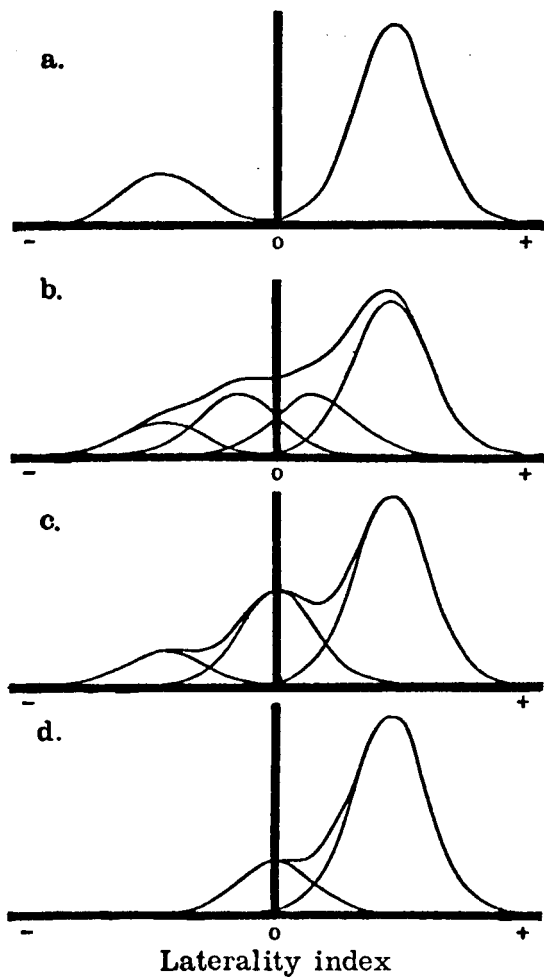
Fig. 2 — *Possible compound distributions which might be of importance in laterality. a - biphasic compound distribution; b - quadriphasic compound distribution (see McManus, 1979); c - triphasic compound distribution, the group with zero mean being regarded as "ambilateral" or "bilateral"; d - biphasic compound distribution as proposed by Annett (1976).*

the data being drawn from a distribution which is compound rather than simple. However we cannot simply compare support values directly since for different models we are using different numbers of free parameters. In an extreme case, if we had n data points and n parameters, each of which represented the value of one data point, we would obtain a "perfect fit", and a support of zero (i.e. a likelihood of one). The conventional way of dealing with this problem (Silvey, 1975) is to consider twice the support

difference as being approximately distributed as chi-squared with k degrees of freedom, where k is the difference in the number of free parameters between the two models being considered. Thus in this case. Model 0 has two parameters and Model IV has five parameters: hence twice the support difference may be considered to be distributed as stet^2 with 3 df. If this chi-squared value is sufficiently large by conventional standards, we may say that Model IV is a significantly better fit than Model 0. It should of course always be a *better* fit (or, at least, as good since the extra parameters are nested inside the earlier parameters); the question is whether it is *significantly* better. If two models have the same number of parameters we may consider their support differences directly, where if d is the difference in support between the two models then e^d gives the relative likelihood of one model with respect to the other. Edwards (1972) recommends that in general a support difference of 2 (i.e. relative likelihood of 7.4) should be a conventional criterion for regarding two models as different in their fit.

By comparing the support values for models 0 to IV we may decide whether it is necessary to postulate two separate distributions, and if so, whether they need be asymmetric for variance and/or means.

In some experiments the observations will themselves have a known error variance, and we may fit models which take account of such factors. Consider the case in which a subject detects r out of t stimuli presented to the right ear, and l out of t stimuli presented to the left ear. We may calculate the Z-transform of the tetrachoric correlation coefficient and use this as a measure of the degree of asymmetry. This value will have an error-variance which is a known function of r and l. Alternatively let a subject make t preference judgements of which r are for an object on the right side and l for an object on the left side. The asymmetry coefficient $2r/t - 1$ will be binomially distributed and have a known error variance, which will be a function of r and t. In such cases we may fit an extended series of models in which ^{2} has two components, $^2_{error}$ due to sampling error, and $^2_{subjects}$ due to differences in asymmetry between subjects, where $^2 = {}^2_{subjects} + {}^2_{error}$. We may thus determine whether it is necessary to postulate true differences between subjects, as opposed to just differences due to sampling errors. That is, can the population be regarded as binomially distributed, or is some other distribution, such as the Lexian, necessary to describe it.

Thus far I have described a method for examining the overall distribution of results from a population, and deciding whether that distribution is best regarded as simple or compound, and if the latter, with what parameters. We may now extend the problem to consider the case where we have two sub-groups, each of which can be regarded in the same way as

the overall population. Specifically, let us consider a group of left-handers and a group of right-handers, and ask whether right- and left-handers differ in their lateralisation. We may divide the question into several components; do they have different means; do they have different variances; and do they have different proportions of types + and $-$?

Let μ_{R+}, μ_{R-}, μ_{L+} and μ_{L-} represent the means of right- and left-handers of types + and $-$. And similarly, let σ_{R+}, σ_{R-}, σ_{L+}, and σ_{L-} represent their standard deviations. Let there be n_R right-handers, and n_L left-handers, totalling N. Let x_i be the score of the ith individual, where if $1 < i < n_R$ then the individual is right-handed, and if $n_{R+1} < i < N$ the individual is left-handed. For a single right-handed individual the support for a particular set of parameters μ_{R+}, μ_{R-}, σ_{R+}, σ_{R-} and π_R is:

$$S_R(x_i) = \log_e [(1 - \pi_R) \cdot N(x_i, \mu_+, \sigma_+) + \pi_R \cdot N(x_i, \mu_-, \sigma_-)]$$

... 7

where π_R is the proportion of right-handers of type $-$. And similarly for left-handers:

$$S_L(x_i) = \log_e [(1 - \pi_L) \cdot N(x_i, \mu_+, \sigma_+) + \pi_L \cdot N(x_i, \mu_-, \sigma_-)]$$

... 8

where π_L is the proportion of left-handers who are of type $-$. And hence the total support for a set of parameters is:

$$S_T(x_i, i = 1, N) = \sum_{i=1, n_R} S_R(x_i) + \sum_{j = n_{R+1}, N} S_L(x_j)$$

... 9

Estimates of the maximum-likelihood values of the parameters may be obtained by similar numerical methods to those described earlier. We may now find the support for different Models.

By setting $\Pi_R = \Pi_L = 0$, $\mu_{R+} = \mu_{L+}$ and $\sigma_{R+} = \sigma_{L+}$ we are fitting Model 0 described earlier, there being a unimodal normal distribution, with no difference between right- and left-handers.

If we set $\pi_R = \pi_L = 0$, $\sigma_{R+} = \sigma_{L+} \cdot$ and $\mu_{R+} \neq \mu_{L+}$ we are fitting an analysis of variance model (in this case, the equivalent of an unpaired t-test), and if we allow $\sigma_{R+} \neq \sigma_{L+}$ we are fitting the model for the equivalent of a t-test in which the distributions have unequal variances. The improvement in fit of these models over model 0, may be tested for significance by comparing their supports.

By fitting a model in which $\pi_R \neq \pi_L \neq 0$, $\mu_{R+} = -\mu_{R-} = \mu_{L+} = -\mu_{L-}$ and $\sigma_{R+} = \sigma_{R-} = \sigma_{L+} = \sigma_{L-}$ we are fitting the equivalent of Model I for population data, in which two groups differ only in their proportion of type minus. And similarly we may allow the means and the variances to differ between groups and types, producing the equivalent of Models II and III. For a fully saturated model, $\pi_R \neq \pi_L$, $\mu_{R+} \neq -\mu_{R-} \neq \mu_{L+} \neq -\mu_{L-}$ and

$\sigma_{R+} \neq \sigma_{R-} \neq \sigma_{L+} \neq \sigma_{L-}$. Each of these models may be tested in the conventional manner to determine whether their fit is better than Model 0, or the standard ANOVA model.

By setting $\pi_R = \pi_L \neq 0$, $\mu_{R+} = \mu_{L+} \neq -\mu_{R-} = -\mu_{L-}$, and $\sigma_{R+} = \sigma_{L+} \neq \sigma_{R-} = \sigma_{L-}$, we are fitting Model IV described earlier.

Having gone thus far in fitting models in which right- and left-handers can be differentiated into + and − sub-types (see Figure 2a), it is apparent that our models can, in principle at least, go one step further and have multiple sub-types. Elsewhere I have proposed a genetic model of cerebral dominance (McManus, 1979) which argues that it is possible that there may be four types of speech dominance labelled as + +, + −, − +, and −−, as assessed by dichotic listening tests; that is, there are individuals with a Strong Right-Ear Advantage (SREA), a Weak Right-Ear Advantage (WREA) and Weak Left-Ear Advantage (WLEA) and a Strong Left-Ear Advantage (SLEA) (see Figure 2b). Within right-handers a family history of sinistrality would change the relative proportions of SREA and WREA and SLEA and WLEA. And it is simple, in principle, to test whether such a model is better supported by the data than is a bimodal normal distribution, or a unimodal normal distribution (although in practice it may be difficult to obtain a sufficient amount of data to test the hypothesis convincingly). Alternatively we may propose a further model in which there are three sub-types +, − and 0, where the 0 group may be described as "bilateral" and have, overall, no lateralisation (see Figure 2c). Or there could be just two sub-types + and 0, as proposed by Annett (1976) (Figure 2d).

Factorial experiments

Consider a more general case in which we have a single asymmetry score from a number of subjects, the subjects being divided into sub-groups on the basis of a number of factors. These factors may, or may not be orthogonal, and hence the experiment need not necessarily be 'balanced' in the sense used in the analysis of variance. Let k be the number of various combinations of factors, and assume, for the sake of demonstration, that there are at least several subjects in each sub-group. Thus if we are considering an experiment in which we are looking at the effects of handedness, sex, and familial sinistrality on laterality, each of which has two levels, then there will be $2 \times 2 \times 2$ sub-groups, and thus k = 8. In the most extreme, fully saturated, case, all the sub-groups will be different from one another, each having its own μ_+, μ_-, σ_+, σ_-, and π. We must therefore fit 5k parameters to our data (see Figure 3). However very many of these parameters will frequently be identical, and we may identify
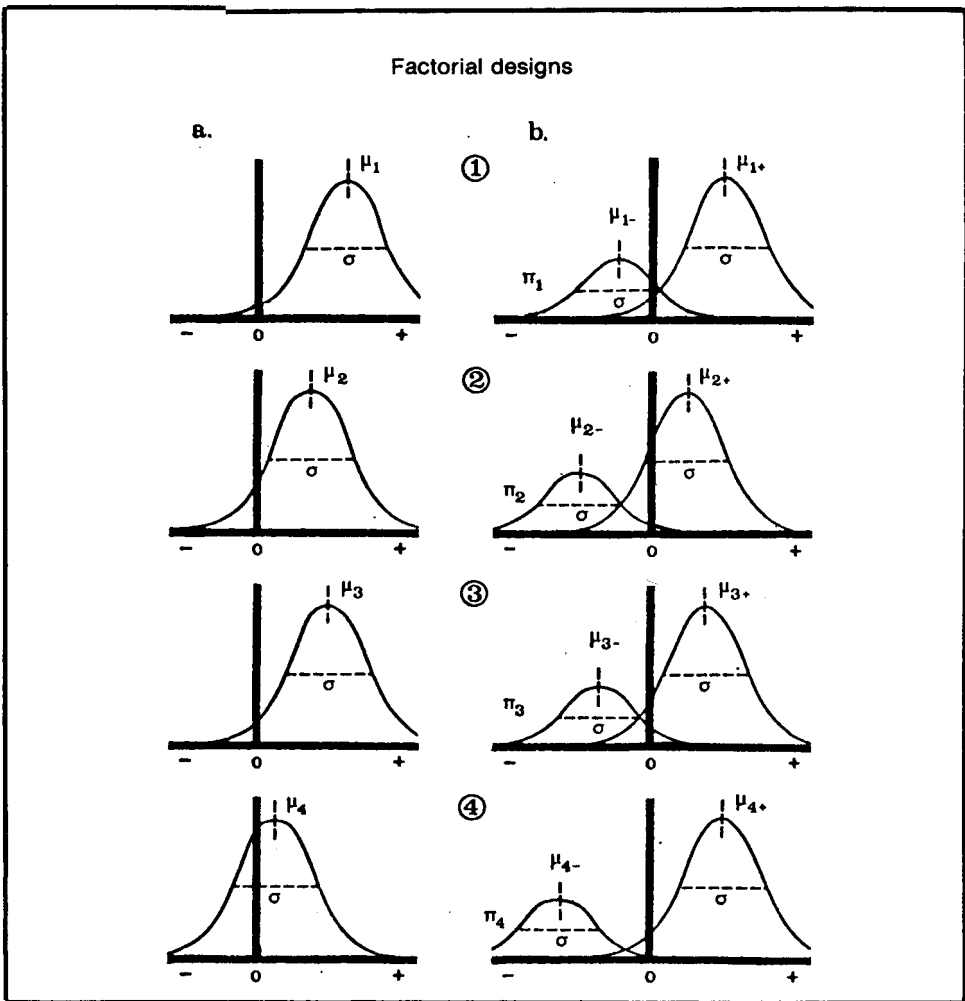
Fig. 3 — *Parameters estimated (a) by a classic analysis of variance, and (b) by a version of the model proposed in this paper (which assumes equal variance in all sub-distributions), there being four sub-groups for each of which there may be different means (both for +ve and −ve sub-types) between groups, and different proportions of type −ve between sub-groups.*

separate effects of the particular parameters. As a global test of *any* effect of handedness, sex or familial sinistrality we can find S_T, the support for the complete model with 5k free parameters. We may also find S_{IV}, the support for a model with five parameters in which all k sub-groups are identical (that is, Model IV, given earlier). $S_T - S_{IV}$ then estimates the improved support obtained by allowing differences between sub-groups. Since support is additive we may partition this support difference in a similar manner to that in which we normally partition variance. A first step would be to fit a model, A, in which for the jth sub-group $\mu_{j+} = -\mu_{j-} =$

μ_{1+}, and $\sigma_{j+} = \sigma_{j-} = \sigma_{1+}$, and for which π_j is allowed to be different within each sub-group, and to find the support, S_A, for such a model. If $S_A - S_{IV}$ is large enough then differences in relative proportions of $+$ and $-$ sub-types are important in differentiating between groups. If $S_T - S_A$ is small enough as well, then all differences between sub-groups can be attributed entirely to differences in the proportions of types $+$ and $-$, rather than to differences in degree or variability of laterality. And, if that is also the case, we may ask whether the differences in proportions are due to, say, just differences between handedness groups, or due to just sex differences, or due to an interaction between the two factors.

It is worth noting that when for all j, $\pi_j = 0$, and $\sigma_{j+} = \sigma_{1+}$, that is all individuals are of type $+$, and the variance of each sub-group is identical, then the analysis reduces to a simple analysis of variance.

Test-retest reliability

Thus far this paper has been concerned only with the randomised blocks design of experiment in which each subject contributes a single number to the data set; that is all experimentally controlled factors are between-subject variables. Frequently, however, one has several observations from an individual subject; that is, the design is of the split-plot form, with one or more within-subject variables. The present method of analysis may be generalised to such data although, as will become apparent, it is rather complex with more than two observations per subject. Before considering the full split-plot design, let us consider the conceptually simpler case in which we wish to assess the split-half or test-retest reliability of a measure of asymmetry. Each subject thus contributes two separate scores. A standard approach to the problem would be to take all of the two sets of data points, which we may call x_i and y_i, where $i = 1, n$ and y_i may be regarded as the 'second' of the two observations, and simply to calculate either a Pearsonian or a Spearman correlation of x_i with y_i. However in terms of our previous analysis, it must be clear that the interpretation of this single number is difficult. Does it mean that the *degree* of asymmetry is measured reliably in both $+$ and $-$ sub-types, or that direction of laterality is reliable, etc.? If there is evidence that x_i or y_i are drawn from compound distributions then the surface obtained by plotting x_i against y_i will not be, as a simple Pearsonian analysis would assume, a simple bivariate normal distribution, but a compound bivariate distribution.

Figure 4 shows computer-generated plots of such probability density surfaces. Figure 4a represents a simple bivariate normal distribution with zero correlation. Figure 4b shows a compound distribution (which we may
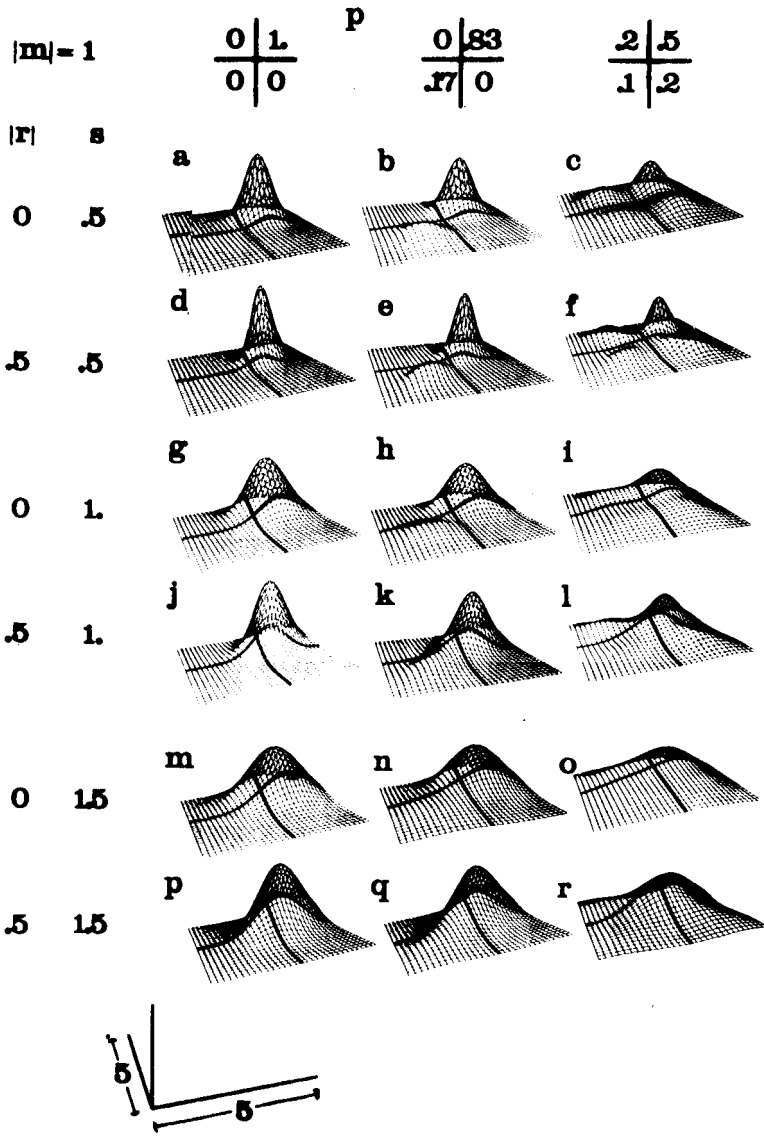
Fig. 4 — *Computer-drawn perspective projections of the probability surfaces of various models. Column 1 shows a single bivariate normal distribution, Column 2 a compound distribution composed of two bivariate normal distributions (in proportions 5:1) and Column 3 a compound distribution composed of four bivariate normal distributions (in proportions 5:2:2:1), the proportions being identical in all rows of the figure. All means are an absolute distance of one unit from the axes. In rows 1 and 2 the standard deviations of each distribution (or sub-distribution) are 0.5, in rows 3 and 4 they are .75, and in rows 5 and 6 they are 1.0. For rows 1, 3 and 5 the correlation within each bivariate normal sub-distribution is zero, while in rows 2, 4 and 6 the correlation is 0.5. Some compound distributions have been scaled by a constant multiplier on the vertical (Z) axis to make their appearances more compatible. The solid cross-lines represent the axes in the x and y dimensions.*

loosely label 'bimodal') in which each 'mode' has zero correlation. Clearly in such a case a simple Pearsonian correlation would shown an artefactually significant correlation. The real question of interest is whether the *degree* of laterality is reliably measured, and thus whether Figure 4e is a better fit of the data than Figure 4b. As the variances of each distribution increase this assessment becomes increasingly difficult by eye, as will be seen by comparing appropriate rows of Figure 4.

If we assume, in the first instance, that all individuals of a particular sub-type will be of the same sub-type on both test and re-test then the surface of x_i against y_i will contain a mode in the $+,+$ quadrant, and another mode in the $-,-$ quadrant. Model IV for a single dependent variable has five separate parameters, μ_+, μ_-, σ_+, σ_-, and π. The bivariate surface for test-retest data will have seven parameters, μ_+, $\mu_-, \sigma_+, \sigma_-, \pi, \rho_+$, and ρ_-, where ρ_+ and ρ_- are the test- retest correlations within a sub-type for each of the sub-types. (See Figure 5).

If we consider the $+$ve sub-type, then the likelihood, L_+ of a pair of observations, x_i and y_i, given a particular set of parameters, is:

$$L_+ (x_i, y_i) = B (x_i, y_i, \mu_+, \sigma_+, \mu_+, \sigma_+, \rho_+)$$

where $B (x, y, \mu_1, \sigma_1, \mu_2, \sigma_2, \rho) = \dfrac{\exp -\left[\dfrac{\dfrac{(x-\mu_1)^2}{\sigma_1^2} - \dfrac{2\rho\,(x-\mu_1)\,(y-\mu_2)}{\sigma_1\,\sigma_2} + \dfrac{(y-\mu_2)^2}{\sigma_2^2}}{2(1-\rho^2)}\right]}{\pi\,\sigma_1\sigma_2\sqrt{(1-\rho^2)}}$

...10

And for the same pair of observations, the likelihood, $L-$, that the data come from sub-type $-$, is:

$$L_- (x_i, y_i) = B (x_i, y_i, \mu_-, \sigma_-, \mu_-, \sigma_-, \rho_-)$$

...11

And hence the likelihood, L, that the observation pair comes from a bivariate distribution with the above parameters is:

$$L (x_i, y_i) = (1 - \pi) \cdot L_+ (x_i, y_i) + \pi \cdot L_- (x_i, y_i)$$

...12

And thus the overall support, S_T, for the particular seven parameters, given all of the data pairs, is:

$$S_T(x_i, y_i, i = 1, n) = \sum_{i=1,n} \log_e [L(x_i, y_i)]$$

...13

We may thence derive the maximum likelihood estimates of μ_+, μ_-, σ_+, σ_-, ρ_+, ρ_-, and π by the numerical method described earlier. It is also, of
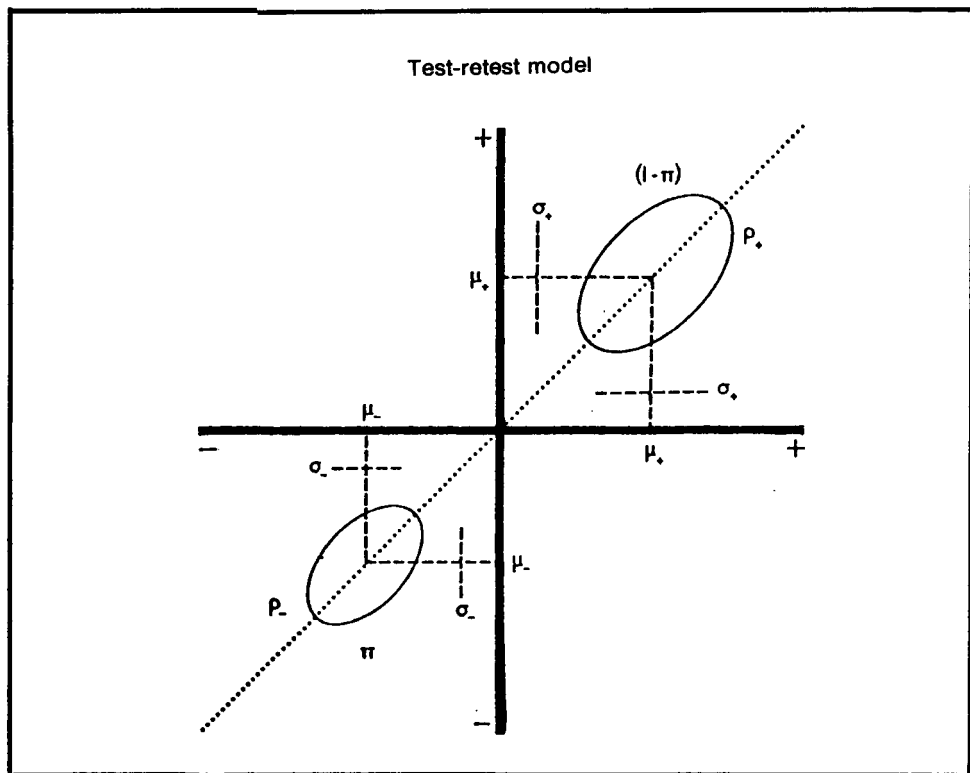
Fig. 5 — *Parameters which may be estimated for a simple test-retest experiment for the assessment of reliability of a measure.*

course, possible to produce simpler models than the full model described in equation 13, and by comparison of these simplified models with the full model it is possible to decide which is the most appropriate model of the data. By analogy with Model I described earlier we may fit a model to test-retest data in which $\mu_+, = -\mu_-, \sigma_+ = \sigma_-, \pi \neq 0$, and $\rho_+ = \rho_-$. If this model is sufficiently well-supported in comparison with the general model, then we may also fit a model which is identical except that $\rho_+ = \rho_- = 0$; that is there is no test-retest reliability of *degree* of handedness within either sub-type, + or −. If the former model does not receive sufficient extra support over the latter model then we may conclude that there is no test-retest correlation of *degree* of laterality, within either the +ve or the −ve sub-type. It is worth noting that even though this might be the case it is still quite likely that a simple Pearsonian correlation of x with y would be significant since that measure is confounded by *direction* of laterality, which is probably reliable (and indeed has been assumed to be so in the present discussion; it may be explicitly tested for by the method to be described in the next section).

A more complex test-retest model may be fitted in which we examine the possibility, ignored thus far, that individuals of type + and − become more lateralised or more variably lateralised between the first and second tests. Let $\mu_{1+}, \sigma_{1+}, \mu_{1-}, \sigma_{1-}$, and π, be the relevant parameters for the first test, and $\mu_{2+}, \sigma_{2+}, \mu_{2-}, \sigma_{2-}$, and π be the parameters for the second test (π being constant since the individuals' sub-type is assumed to be fixed), then by analogy with equations 10 and 11:

$$L_+ (x_i, y_i) = B (x_i, y_i, \mu_{1+}, \sigma_{1+}, \mu_{2+}, \sigma_{2+}, \rho_+)$$

...14

$$L_- (x_i, y_i) = B (x_i, y_i, \mu_{1-}, \sigma_{1-}, \mu_{2-}, \sigma_{2-}, \rho_-)$$

...15

and hence we may use equations 12 and 13 to derive S_T, the overall support for the model, and we may then derive the maximum likelihood estimates for the relevant parameters.

By comparing the support derived from equations 14 and 15, with that derived from equations 10 and 11, we may decide whether there is adequate evidence for differences in lateralisation between the test and the re-test.

In principle the above method could be extended to deal not only with population data but also to factorial experiments, but discussion of this will be deferred until the full split-plot analysis has been described.

Split-Plot designs

Consider an experiment in which a number of subjects are tested for their degree of lateralisation on both a verbal and a visuo-spatial task: that is, task type is a within subject variable. We could analyse the data from such an experiment by using the more complex form of the test-retest model, in which we allowed the possibility that the tasks (tests) might differ in their means and variances for both the + ve and −ve sub-types. However this model fails to account for the possibility (which can reasonably be excluded *a priori* in the test-retest case) that some individuals might be of the + ve sub-type on one task and of the −ve sub-type on the other task. (By + and − I here mean "dominant" and "non-dominant", as applied to the population overall, such that the + type would be left-hemisphere dominant for the verbal task, and right-hemisphere dominant for the visuo-spatial task).

In stating the problem in this form it is immediately apparent that we must also allow for a further possibility; that the overall proportion of the − sub-type might be different for the two tasks.

Let us label the tasks as A and B. Let $\mu_{A+}, \mu_{A-}, \mu_{B+}, \mu_{B-}, \sigma_{A+}, \sigma_{A-}, \sigma_{B+}$, and σ_{B-} be the appropriate means and standard deviations of the four possible sub-types. Let π_A and π_B be the proportions of sub-type − for the A and B tasks. Further let π_{AB} be the proportion of the population who are of

sub-type – on *both* tasks. Using this nomenclature it is apparent that if π_A = π_B = π_{AB} then there is a perfect correlation between the tasks in their direction of lateralisation, and if $\pi_{AB} = \pi_A \cdot \pi_B$ then there is zero correlation between the two tasks in their *direction* of lateralisation. It is trivially necessary that if $\pi_A \neq \pi_B$ then there must be at least some individuals of either sub-type $-/+$ or sub-type $+/-$. In general it will be true that $\pi_A + \pi_B \geqslant \pi_{AB} \geqslant \pi_{A \cdot B}$. If $\pi_{AB} < \pi_A \cdot \pi_B$ then there is crosslaterality; that is, the majority of individuals are oppositely lateralised for the two tasks; and in general this will merely mean that the conventional description of $+$ and $-$ has been reversed, since the "minority" form will be in the majority.

The value of π_{AB} relative to π_A and π_B is therefore a measure of consistency in *direction* of asymmetry. As before, in the test-retest case, we may also have consistency in *degree* of laterality, which in the split-plot case is better described as a *correlation* of degree of lateralisation in the two tasks. There are four possible sub-types, $+/+$, $+/-$, $-/+$ and $-/-$. For each sub-type there may be a separate correlation between the two tasks, and we may label these correlations as ρ_{++}, ρ_{+-}, ρ_{-+} and ρ_{--}. (See Figure 4c, f, i, l, m and q for computer plots of typical distributions). In general it will usually be assumed in the first instance that all of these correlations are equal although that assumption is one capable of empirical test. For groups $-/+$ and $+/-$, the correlation will, of course, be $-\rho_{++}$ if symmetry is to be retained. It is possible to imagine cases in which individuals of sub-type $+/+$ and $-/-$ would show a strong correlation in their degree of lateralisation, but sub-types $+/-$ and $-/+$ would show no such correlation. The four correlations are therefore independent of one another. We may also extend the parameters described earlier to account for the possibility that all of the four sub-type combinations differ in their means and variances on each of the tests for both $+$ and $-$ sub-types. For the $+/+$ sub-type combination we must therefore specify five parameters, μ_{A++}, μ_{B++}, σ_{A++}, σ_{B++} and ρ_{++}; and similarly for the other three sub-type combinations. As well as these twenty parameters, we must also specify π_A, π_B, π_{AB} in order to completely specify the model; a total of twenty-three parameters (see Figure 6). If we wished to ignore the possibility of interactions between the tests in their degree of lateralisations we could reduce these parameters to 15, since $\mu_{A++} = \mu_{A+} = -\mu_{A+-}$, etc.

Consider a pair of observations, x_i, y_i from a single subject. For the full model, with 23 parameters, we may derive the likelihood, L_{++} that this individual is from the sub-type combination $+/+$ with the particular parameters:

$$L_{++} (x_i, y_i) = B (x_i, y_i, \mu_{A++}, \sigma_{A++}, \mu_{B++}, \sigma_{B++}, \rho_{++})$$
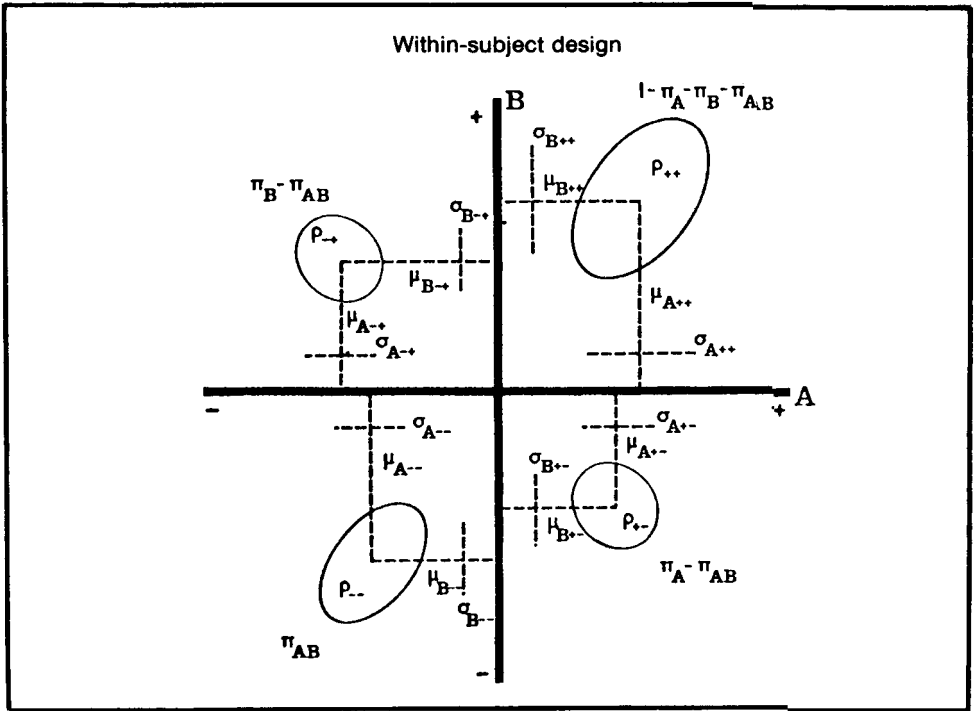
...16

Fig. 6 — *Possible parameters which may be fitted for a saturated model in which each subject carries out two separate tasks and obtains two separate laterality scores.*

and similarly for L_{+-}, L_{-+}, and L_{--}. We may then derive L, the likelihood of the total set of parameters given the single pair of observations x_i and y_i.

$$L(x_i, y_i) = \pi_{AB} \cdot L_{--}(x_i, y_i) + (\pi_B - \pi_{AB}) \cdot L_{+-}(x_i, y_i)$$
$$+ (\pi_A - \pi_{AB}) \cdot L_{-+}(x_i, y_i) + (1 - \pi_A - \pi_B + \pi_{AB}) \cdot L_{++}(x_i, y_i)$$
$$...17$$

And hence S_T, the support for the particular set of parameters is:

$$S_T(x_i, y_i; i = 1, n) = \sum_{i = 1, n} \log_e (L[x_i, y_i])$$

$$...18$$

Equations 16... 18 may be solved by means of the hill-climbing technique described earlier, in order to obtain the maximum likelihood estimates of the 23 parameters. And, as for the earlier models, so it is possible to find the support for reduced versions of the model with reduced numbers of free parameters in order to determine whether there is an adequate increase in support for the more complex models over the simpler ones.

It should by now be obvious that the split-plot model may be further expanded to fit more complex experimental designs. In equations 16... 18 only two separate tasks have been considered, but in principle more complex models involving three or more tasks may be analysed by using the multivariate normal distribution instead of the bivariate normal distribution, and increasing the number of sub-type combinations and parameters appropriately. In practice the number of parameters rises exponentially with the number of separate tasks, and the equations rapidly become impossibly complex unless the models are constrained fairly tightly. A second possibility for expansion of the split-plot model is to consider the analysis of experiments in which there are both within-subject variables and also between-subject variables. Consider an experiment in which subjects are divided on the basis of handedness, sex, and familial sinistrality, and that all of the eight possible sub-groups are represented. For each sub-group we may fit a model similar to that in equations 16... 18, and it is possible that the parameters for each sub-group are different from those for all other sub-groups. In such a case there would therefore be a total of $8 \times 23 = 184$ free parameters to be fitted in the complete model. Clearly such models are also unwieldly unless some constraints are put upon the parameters.

Example 1

As an illustration of the method, the randomly generated data described earlier will be analysed. For those who might wish to re-work this demonstration the values for the "right-handed" group are
$-1.235, -.876, 1.379, 1.116, .829, .486, .650, 1.241, 1.491, 1.824, .858, .627,$
$1.152, 1.568, 1.133, .734, .253, .985, 1.574$ and 1.207, whilst those for the "left-handed" group are $-1.666, -1.354, -1.077, -1.700, -2.343, -1.006,$
$-.565, 1.207, 1.352, 1.760, .077, .867, 1.158, 1.046, .874, 1.669, .619, 1.138,$
$.770$ and 1.052. The numbers were drawn from the distributions already described, using the uniform random-number generator of a Commodore PET 2001 microcomputer, and the normal distributions being derived from this by the algorithm described by Knuth (1969).

The simplest model which may be fitted to these data is that in which both groups are drawn from the same normal distribution (Model A, Table I). If we now fit a model in which the two groups have different means, but identical variances (Model B) then we obtain an improved support, the differences in support being equivalent to a chi² value of 3.867 (p = .0492), a value which is similar in significance to the value of t of 1.917 (p = .0628) described earlier (and this is, indeed, an analogous test). If we fit a symmetric bimodal distribution (Model C) (type I described earlier) then there is a much improved fit over the Classic ANOVA model

TABLE I
Equations

| Model	Description	Parameters	Group R	Group L	Support	N. Parameters	Maximum likelihood estimates of parameters
A	Single Normal	μ, σ	$y = N(\mu, \sigma)$	$y = N(\mu, \sigma)$	59.842	2	$\mu = .522\ \sigma = 1.080$
B	Classic ANOVA	μ_R, μ_L, σ	$y = N(\mu_R, \sigma)$	$y = N(\mu_L, \sigma)$	57.908	3	$\mu_R = .849\ \mu_L = .1938$ $\sigma = 1.029$
C	Symmetric Normal	μ, σ, π	$y = (1-\pi)N(\mu,\sigma) + \pi N(-\mu,\sigma)$	$y = (1-\pi)N(\mu,\sigma) + \pi N(-\mu,\sigma)$	45.793	3	$\mu = 1.112\ \sigma = .450\ \pi = .228$
D	Symmetric Normal Different proportions	$\mu, \sigma, \pi_R, \pi_L$	$y = (1-\pi_R)N(\mu,\sigma) + \pi_R N(-\mu,\sigma)$	$y = (1-\pi_L)N(\mu,\sigma) + \pi_L N(-\mu,\sigma)$	43.839	4	$\mu = 1.112\ \sigma = .450\ \mu_R = .100$ $\pi_L = .359$
E	Symmetric Normal Different means and proportions	$\mu_{R+}, \mu_{R-}, \mu_{L+}, \mu_{L-}, \sigma, \pi_R, \pi_L$	$y = (1-\pi_R)N(\mu_{R+}, \sigma) + \pi_R N(-\mu_{R-}, \sigma)$	$y = (1-\pi_L)N(\mu_{L+}, \sigma) + \pi_L N(-\mu_{L-}, \sigma)$	42.359	7	$\mu_{R+} = 1.062\ \mu_{R-} = 1.050$ $\mu_{L+} = 1.045\ \mu_{L-} = 1.384$ $\sigma = .431\ \pi_R = .100\ \pi_L = .351$
F	Symmetric Normal Different means, variances and proportions	$\mu_{R+}, \mu_{R-}, \mu_{L+},$ $\mu_{L-}, \sigma_{R+}, \sigma_{R-},$ $\sigma_{L+}, \sigma_{L-}, \pi_L, \pi_R$	$y = (1-\pi_R)N(\mu_{R+}, \sigma_{R+}) + \pi_R N(-\mu_{R-}, \sigma_{R-})$	$y = (1-\pi_L)N(\mu_{L+}, \sigma_{L+}) + \pi_L N(-\mu_{L-}, \sigma_{L-})$	40.686	10	$\mu_{R+} = 1.061\ \mu_{R-} = 1.055$ $\mu_{L+} = 1.112\ \mu_{L-} = 1.219$ $\sigma_{R+} = .408\ \sigma_{R-} = .179$ $\sigma_{L+} = .347\ \sigma_{L-} = .694$ $\pi_R = .099\ \pi_L = .394$
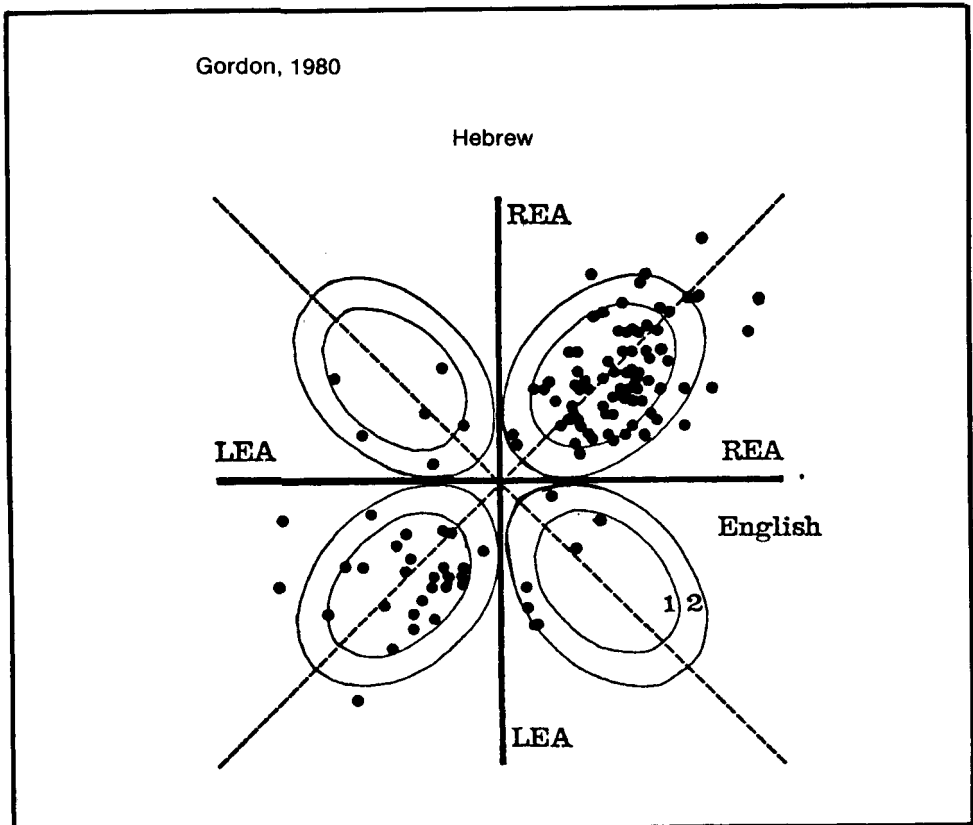
Fig. 7 — *The data of Gordon (1980; his Figure 2) replotted. Each point represents a single subject (n = 120). The abscissa and ordinate are plotted on the same non-linear scale as that of Gordon, and the ordinate has been stretched slightly so that the 45° line represents equal scores on the two tests. The ellipses represent 1 and 2 support units from the proposed modes, the parameters being those of model E.*

(Model B), the support difference being 12.12, being equivalent to a likelihood ratio of 182608.5 in favour of Model C. If we fit a symmetric bimodal distribution in which the proportion of type minus is different in the two sub-groups (Model D) then this is better supported than Model C, suggesting that the two groups differ in their proportion of type minus ($chi_1^2 = 3.907, p = .0481$). Model D fits the data very much better than the Classic ANOVA model $chi_1^2 = 28.137, p < 0.001$).

More sophisticated versions of the bimodal model, in which the means and proportions (Model E), or means, variances and proportions (Model F) differ between groups do not fit the data better than Model D $chi_3^2 = 2.96$ and $chi_6^2 = 6.306$ respectively).

The analysis therefore shows clearly that the distributions are compound, suggests that the proportions of type minus differ between the

groups, but finds no evidence for the sub-groups differing in their means and variances.

Example 2

The second example concerns repeated measures on a number of subjects. Gordon (1980) reported data of some considerable theoretical interest on the lateralisation of bilinguals. A large number of right-handed Israelis who were bilingual for English and Hebrew were tested on a dichotic listening task, the words being presented in each of the languages, and laterality indices being scored separately for the two languages. Figure 7, redrawn from Figure 2 of Gordon (1980), shows the individual laterality scores of 120 subjects in both Hebrew and English. (The axes are linearly rated to the axes in Gordon, 1980, *as drawn*; however Gordon, 1980, himself had plotted these values non-linearly — see his text for details). Gordon himself concludes from an examination of the results that "An even stronger indication of the lateral dominance of the two languages is the extremely high correlation between the Hebrew (index of laterality) and the English (index of laterality) (r = .78 for all subjects counted together). Thus, a high lateralisation in one language was coupled with a high lateralisation in the other language. Or conversely, a small or reversed (left ear) lateralisation in one language was usually accompanied by a similar reversal in the second language". (p 259).

Scrutiny of the data of Figure 7 suggest that there are a number of further questions that we may ask about the results apart from the correlation. There appears to be clear visual evidence for a bimodal distribution (i.e. some individuals appear to show atypical lateralisation for both two languages).

Furthermore it appears that the correlation *within* each of the modes is relatively low and may perhaps not be different from zero. Finally we may ask whether the degree of lateralisation of the modes is the same in the two languages.

Table II summarises the results of the fitting of a series of models. Models are fitted in which there are one, two or four modes, and for which the correlation within a mode is zero or non-zero. Models are also fitted in which the means and standard deviations within modes are allowed to differ for the two languages. The model assumed by Gordon is model B with $r \neq 0$. It may be seen that model D and model F, with $r \neq 0$, are significantly better fits (chi_1^2 = 48.0, p < 0.001; chi_3^2 = 66.406, p < 0.001 respectively). Furthermore model F with $r \neq 0$ is a significantly better fit than model D with $r \neq 0$ (chi_2^2 = 20.706, p < 0.001). Model F with $r \neq 0$ is also significantly better at fitting the data than model F with $r \neq 0$ (chi_1^2 = 16.256, p < 0.001). Model E with $r \neq 0$ is however not a significant worse fit than model F with $r \neq 0$ (chi_2^2 = 4.512, n. s.); there is therefore no need

TABLE II

The Support and Some of the Estimated Parameters for Unimodal, Bimodal and Quadrimodal Distributions Fitted to the Data of Gordon (1980)

	N (Parameters)	Parameters	Support	Within-group correlation ρ	$\pi(+\,+)$	$\pi(+\,-)$	$\pi(-\,+)$	$\pi(-\,-)$	Support for equivalent model with $\rho = 0$
Unimodal models									
A	3	μ, σ, ρ	−1028.060	.7679	1.	0.	0.	0.	–
B	5	$\mu_E, \sigma_E, \mu_H, \sigma_H, \rho$	−1026.046	.7743	1.	0.	0.	0.	−1080.976
Bimodal models									
C	4	$\mu, \sigma \rho, \pi$	−1005.452	.2587	.7134	0.	0.	.2866	–
D	6	$\mu_E, \sigma_E \mu_H, \sigma_H, \rho, \pi_{--}$	−1002.674	.3499	.7266	0.	0.	.2774	−1008.275
Quadrimodal models									
E	6	$\mu, \sigma, \rho, \pi_{+-}, \pi_{-+}, \pi_{--},$	− 995.099	.3511	.6818	.0328	.0389	.2463	–
F	8	$\mu_E, \sigma_E \mu_H, \sigma_H, \rho, \pi_{+-}, \pi_{-+}, \pi_{--}$	− 992.843	.3644	.6816	.0335	.0389	.2460	−1000.971

to make the assumption of different means and variances between languages.

We may summarise the analysis as follows: there is clear evidence of a *quadrimodal* distribution, with a small proportion of individuals (7.2%) showing opposite directions of lateralisation for the two languages. Within modes, the degree and variability of lateralisation of the two languages is the same, and there is a smallish correlation (0.35) between *the degrees of lateralisation* in the two languages.

In so far as these conclusions differ from those of Gordon (1980), the utility of the method is demonstrated. The conclusions also have substantial implication for the neuropsychology of bilingualism.

It should of course be clear that a considerable number of further models could still be fitted in which some or all of the remaining equality constraints are relaxed in order to see if a further improvement in fit may be obtained. Since however the purposes of the example have been sufficiently served, this will not be carried out here.

Discussion

"The Analysis of Variance" is perhaps poorly named insofar as it is strictly "*An* analysis of variance", and is only one of a whole set of such analyses, one which, by making a number of assumptions, can be fitted without too much conceptual difficulty to agricultural experiments (its original purpose), and which can very often be successfully applied to psychology experiments. But that it *is* a model must not be forgotten; thus Lovie (1979) quotes the almost despairing comment of Draper and Smith (1966): "The question, 'What model are you considering?' is often met with, 'I am not considering one — I am using Analysis of Variance'."

The implicit models of psychology and of agriculture are not the same, and in some cases *the* analysis of variance must be regarded as potentially misleading. One such case is probably that of laterality; whether it is indeed such a case must be empirically determined by the analysis of actual data.

In the models outlined above I have suggested that direction and degree of laterality must be carefully distinguished, and that in general the basic model for asymmetry is one in which a minority of individuals are not *less lateralised* than the majority but are *oppositely lateralised,* such that the mean degree of laterality seems to be less. There are obvious biological precedents for such a theoretical position. Thus, in the majority of the population the heart is on the left-hand side. In a population of individuals with bronchiecstasis a lesser proportion of individuals will have their heart on the left-hand side, due to the occurrence of Karta-

gener's Syndrome, an association of *situs inversus totalis* and chronic bronchiectasis. But we would not describe the bronchiectatics as being *less* lateralised, in the sense that on average their hearts were nearer to the midline than in a control group, but rather we would describe the bronchiectatic population as having a higher proportion of an atypical sub-type who are exact mirror-images of the normal individuals. The mean absolute distance of the heart from the midline in the two populations would therefore be identical. Such a model is in the first instance, surely the most reasonable for functional lateralisation in the brain, so that the simplest explanation of differences in mean lateralisation coefficient of different groups is different proportions of a minority sub-group, rather than difference in degree of lateralisation of those individuals. Differences in means and variances might occur as well, but the hypothesis cannot reasonably be tested in the absence of testing for differences in direction of lateralisation.

Since the analysis of variance was developed for analysing agricultural data it is worth considering an agricultural example in which, as in the case of laterality, ANOVA would give mis-leading results. Consider a chemical which applied to a plant made some individual plants grow away from light instead of towards it. Let the degree of negative phototropism of a plant be recorded as a negative quantity. In such a case the mean degree of phototropism would confound direction of phototropism and amount of phototropism, and an analysis of variance would be unable to distinguish between the chemical altering the *degree* of phototropism or its *direction.*

The Analysis of Variance in its classic form is a model for experimental observations, as also are the models that I have outlined above. It is possible that *the* analysis of variance is indeed an adequate model of laterality data. This possibility must however be explicitly tested for by fitting firstly the bimodal models described above, and then fitting the standard analysis of variance model (equivalent to fitting a bimodal model with all proportions of the sub-type set to zero and all variances set to be equal) and determining whether the support difference between the two is small enough to conclude that the simpler, classical model may be used.

Whether models of the present form or of the Classic ANOVA type are a better fit to laterality data, particularly those from dichotic listening and tachistoscopic experiments is an empirical question. Whichever model is indeed correct, we will still be presented with theoretical difficulties. If the present models are indeed better then we are unable to interpret the vast majority of experiments that have already been carried out, and their data will need re-analysis. Alternatively, if the classic ANOVA is an appropriate model for, say, differences between right- and left-handers on a

dichotic task, then we have to explain how it can be that right- and left-handers show differences in direction of lateralisation when we interpret the results of unilateral lesions producing aphasia, but show only differences of *degree* in dichotic tasks; the implication would be that the methods are assessing different underlying functional asymmetries.

The vast number of free parameters in say, a split-plot design with factorial between-subject variables might deter many researchers from wishing to use such models. It is not intended that in most cases such complex models should necessarily be fitted. And on theoretical grounds it seems unlikely (and perhaps essential if we are to ever comprehend the phenomena) that *all* the parameters will truly be different and independent of all the others. We may assume, in the first instance at least, models which are relatively simple, for otherwise we will never progress at all. But all models are merely *models,* and hence first approximations to the actual situation. Nevertheless we must not fit overly simplistic models which do not take account of the biological nature and the physical peculiarities of asymmetries. Researchers are therefore welcome to fit much reduced versions of the models given earlier. The crucial thing is that it is explicitly stated that many simplifying assumptions have been made in fitting such models, and that it is recognised that these assumptions may be tested empirically by fitting more complex models.

In psychology theoretical speculation often pushes far ahead of the empirical support for such speculations. Good examples of this are the recent spates of literature in which it is suggested that the degree of lateralisation is different in males and females, or in left- and right-handers, or that the inter-relations of lateralisation for verbal and visuo-spatial tasks are different in males and females, or in good and poor readers. In all of the above examples a conventional analysis of variance is the theoretical model used to analyse the data. When a group of individuals is described as being "more variable in their laterality" or "less well lateralised" than another group, do we mean that within sub-types the mean or the variance is different, or that the proportion of the sub-types is different between the groups? The present method of analysis allows specific answers to such questions.

The method does however raise some technical problems. The calculations for the maximum likelihood values are not quick, even with a modern computer, the values converging very slowly using the quasi-Newtonian method, and the time taken being exponentially related to the number of free parameters. Furthermore it is possible, unless fairly strict control is exercised over the hill-climbing program, to arrive at local maxima, rather than the true maximum value. More serious as a problem is that only raw data from previous experiments may be re-analysed in this fashion. In particular the summary of a set of data provided by a con-

ventional analysis of variance table is not sufficient to allow re-analysis of the data in the present manner, since the present method represents a generalisation of the analysis of variance, and hence information is necessarily lost by presenting it in the form of an analysis of variance. It is also worth noting that the presentation of any form of summary table based upon the present method would not be a satisfactory solution to the fitting of future, perhaps more complex, models, since the present models are, inevitably, sub-sets of yet more complex and possible models. The only solution, if a re-interpretation of data is to be possible, is for all papers to contain, preferably as an appendix in a small type-face, the raw data on which the analysis is based. Alternatively data from experiments should be deposited in an archive of machine-readable data of the form that has already been developed by the Social Science Research Council (see McManus, 1976). Only in these ways can the results of experiments be prevented from being fossilised or petrified in the form of the theoretical models then extant.

ABSTRACT

It is suggested that many analyses of laterality are fundamentally confused in that they fail to distinguish between differences in *degree* of laterality and differences in *direction* of laterality, and that direction and degree of laterality have different biological and psychological interpretations. Conventional statistical tests, such as analysis of variance or non-parametric procedures, are unable to differentiate the two measures, and necessarily produce uninterpretable results. A maximum likelihood method is described which can discriminate between direction and degree, and its applications to factorial, test-retest, and repeated measure designs is explained. Two worked examples are also given, one based on hypothetical data, the other on actual data.

REFERENCES

ADBY, P.R., and DEMPSTER, M.A.H. *Introduction to Optimisation Methods.* London: Chapman and Hall, 1974.

ANNETT, M. Handedness and the cerebral representation of speech. *Annals of Human Biology 3:* 317-328, 1976.

BHATTACHARYA, C.G. A simple method of resolution of a distribution into Gaussian components. *Biometrics 23:* 115-135, 1967.

COHEN, A.C. Estimation in mixtures of two normal distributions. *Technometrics 9:* 15-28, 1967.

DRAPER, N.R., and SMITH, H. *Applied Regression Analysis.* New York: Wiley, 1966.

DUNLAP, J.W. Recent advances in statistical theory and applications, *American Journal of Psychology 51:* 558-571, 1938.

EDWARDS, A.W.F. *Likelihood,* Cambridge: Cambridge University Press, 1972.

GORDON, H.W. Cerebral organisation in bilinguals: I. Lateralisation. *Brain and Language* 9: 255-268, 1980.

HASSELBLAD, V. Estimation of parameters for a mixture of normal distributions. *Technometrics 8*: 431-444, 1966.

KNUTH, D.E. *The Art of Computer Programming. Vol. 2: Semi-numerical algorithms,* Reading, Massachusetts: Addison-Wesley, 1969.

LOVIE, A.D. The analysis of variance in experimental psychology: 1934-1945. *British Journal of Mathematical and Statistical Psychology 32:* 151-178, 1979.

MCMANUS, I.C. Archive for research data (letter). *Lancet 1:* 1188, 1976.

MCMANUS, I.C. *Determinants of Laterality in Man,* University of Cambridge: unpublished Ph. D. thesis, 1979.

NUMERICAL ALGORITHMS GROUP (1981) *NAG Fortran Library Manual, Mark 8,* vol. 3; Numerical Algorithms Group, Oxford.

PEARSON, K. Contributions to the mathematical theory of evolution. *Philosophical Transactions of the Royal Society 185:* 71-110, 1894.

SILVEY, S.D. *Statistical Inference.* London: Chapman and Hall, 1975.

MCMANUS, I.C. *Ph. D, Academic Dept of Psychiatry,* St. Mary's Hospital (Harrow Road) Harrow Road, London W9 3RL

McManus, I.C. Ph. D, Academic Dept of Psychiatry, St. Mary's Hospital (Harrow Road) Harrow Road, London W9 3RL